International Journal of Applied Technology & Leadership
ISSN 2720-5215
Volume 1, Issue 6, July 2022
ijatl@org

SAFETY, BLACK SWANS AND SIFS

McKell W Sanderson

Capitol Technology University, USA

Justin Thygerson

Capitol Technology University, USA

Linda F. Martin

Embry Riddle Aeronautical, USA

Introduction

There are many challenges to the models that are used to forecast success. These models are utilized in a variety of industries, and are not unique to risk management, safety, or any other profession. However, at the end of the day, a model is just a model. It will never exactly capture reality, it will never allow for completely accurate forecasts. The authors are not saying that models are useless, simply reminding the reader that models have many flaws.

One of the major flaws of models is the inability to predict rare events (Taleb, 2010) or so called Black Swan events. The concept of a Black Swan has become a mainstream argument in the practice of risk management. This concept was popularized by Taleb (Taleb, 2010). Taleb argues that a Black Swan is an event that carries the 3 following characteristics:

- An Outlier – it is not detectable easily (if at all) via conventional methods
- High impact – It seems to "punch above its weight" in terms of effects
- It is easily explained *after it has already occurred* through hindsight bias and narrative fallacies

There are several examples given by Taleb as Black Swans, to name a few:

- The events of September 11, 2001
- The rise of Google
- The rise of Hitler to power

Terje Aven also describes Black Swans as rare events, but breaks them into three categories:

- Events entirely outside understanding or expectation (unknown – unknowns)
- Events outside of 'conventional wisdom' but known to some specialists (unknown – knowns)

- And events known, but believed to be negligible in impact or not probable to occur (Aven, 2015)

Both of the above definitions, and there are many others as well, focus on events that are unlikely/improbable to occur, but occur none the less. Black Swans create problems in the risk management setting, because the risk score associated with them is usually not computed, or it is artificially low (due to outlier status, they are by definition not likely to occur). Because of their extreme impact, these events tend to be remembered, and they are often remembered as a time when the risk management process failed.

Aven points out that the "unexpectedness" of the event is subject to the perception of the person experiencing the event (Aven, 2015). For example, if we consider the act of driving, we expect there to be accidents. Seeing others experience an accident is an expectation for many motorists. However, being involved in an accident is usually unexpected. Motorists often believe (rightly so in many cases) that if they operate safely, they will not experience a motor vehicle crash. Getting into a motor vehicle crash is often a black swan for at least one of the parties involved in the crash.

This concept of perception becomes important when discussing "safety culture" for a given company. Safety culture is often measured via perception surveys sent to multiple groupings within an organization and comparing the results to one another. Often, one group feels confident in the safety efforts and culture of the organization, while others feel less so. This is often attributed to lack of communication (Chute & Wiener, 1995). In terms of safety, this means that a company could experience an event that is a Black Swan to some, and not to others. Taleb illustrates this idea with what he calls a "Turkey Problem." The Turkey Problem is:

> "Consider a turkey that is fed every day. Every single feeding will firm up the bird's belief that it is a general rule of life to be fed every day by friendly members of the human race 'Looking out for its best interests' as a politician would say. On the afternoon of the Wednesday before Thanksgiving, something *unexpected* will happen to the turkey. It will incur a revision of belief." (Taleb, 2010)

The event is certainly unexpected for the turkey, and highly impactful. It is much less so for the farmer. These challenges, among others, make the identification of Black Swans difficult, and what is hard to identify is even harder to measure.

In the safety profession, black swan events, often go by a different name. They are severe injuries and fatalities, or SIFs (sometimes these are noted as significant injuries and fatalities). SIFs fit the black swan model, because they are not easily detectible with conventional methods (otherwise they would be stopped), they are high impact, and they are often explained away in hindsight bias laden investigations.

Current practices in the safety industry may lead to an increase in black swan events, rather than decreasing them. This article will explore some of these practices, identify the pitfalls associated with them, and offer suggestions for how they may be improved.

The shaky foundations

Foundational courses in safety science often start with Heinrich's triangle/pyramid (see Figure 1). Herbert William Heinrich was a safety pioneer in the 1930's. Heinrich, reportedly, looked at thousands of data on industrial accidents and wrote the book "Industrial Accident Prevention: A Scientific Approach" in 1931. The book became foundational for the industrial safety, loss control industries. Heinrich's pyramid theory, as it would later be known, described a correlation between serious injuries and unsafe behaviors. Heinrich's pyramid showed 300 unsafe actions led to approximately 1 serious injury (Heinrich, 1931).

Heinrich's pyramid leaves a lot open for debate. Because his data are not available, it is extremely difficult to completely validate or refute his position. Researchers have long suspected that "the principle of a fixed Heinrich ratio has a dubious logical foundation." (Gallivan et al., 2008). Fred Manuele, wrote extensively on Heinrich's pyramid and theories, and determined them to be a myth (F. Manuele, 2003), and Heinrich's ratio has been invalidated by others as well (Gallivan et al., 2008). Logically, this makes sense, preventing papercuts, even 300 of them, does not lead to a reduction in fatal work. The two variables are not related. Gallivan proposed a thought experiment to further solidify this idea.

> "A version of Russian roulette is played with a revolver that has six chambers. One chamber is loaded with a live cartridge, 2 are loaded with blank cartridges and the other three are left empty. The chambers are spun several times, the player places the gun against his head and pulls the trigger. Analysis of outcomes from this dangerous game show a remarkably consistent pattern. The relative frequencies of deaths, powder burn injuries and no injuries were in the ratio 1 : 2 : 3. In an effort to make the game safer (equivalent to a patient safety intervention), a new loading pattern was adopted with one live cartridge, one blank and four chambers left empty. This halved the proportion of powder burn injuries, but had no affect on the proportion of deaths, in contravention of a fixed Heinrich ratio." (Gallivan et al., 2008)

Even though it has been repeatedly discredited, Heinrich's pyramid, is still considered foundational. It is taught, and retaught, in nearly every University Course with an "Intro to Occupational Safety" title. Supporters of the pyramid often refer to more modern versions of Heinrich's pyramid such as Frank Bird's study of the ratio (Bird & Germain, 1996). Conoco-Phillips further restudied the ratio and published their own pyramid (Canfora et al., 2018). However, these updated pyramids, while providing different ratios, do not address the fundamental flaws of Henrich's pyramid.

A quick search of the "safety pyramid" on Google Scholar will reveal that this debate is not over. For every article discrediting the pyramids, you can find one that is affirming them. This phenomenon may be attributable to the lack of repeatability in Social Sciences research (Sanderson & Martin, 2022). However, many of the articles affirming them are not defending the specific ratios, but rather are defending the idea that small injuries left untreated lead to much larger injuries down the road.

Heinrich's pyramid, led directly to the development of Heinrich's Domino theory. Domino Theory summarized states that five sequential actions occur resulting in an injury (or incident). These factors are:

1. Ancestry and social environment
2. Fault of the person
3. Unsafe Act
4. Accident
5. Injury (Shahab Hosseinian & Jabbarani Torghabeh, 2012)

The Domino Theory was first published in 1931, and was the dominant theory for nearly 50 years, though others were certainly present during the time. Heinrich's model oversimplifies human behavior in the cause of workplace accidents, and seems to focus on blaming the individual for the injury. Look at the language used in the theory, terms such as: "ancestry," "Fault of the person" and "Unsafe act" are all examples of loaded language. Utilizing his own theory, Heinrich concluded that 88 percent of all workplace accidents were attributable to "unsafe acts" by humans (Shahab Hosseinian & Jabbarani Torghabeh, 2012).

And similarly to the pyramids, rather than simply identifying a mistake and moving away from it, researches would spend years trying to "fix" the Domino Theory. Frank Bird (the same Frank Bird from above) changed the names of the Dominos but kept the basic idea of the Domino Theory in 1974. His Domino Theory (developed along with Loftus in 1974), titled the Updated Domino Theory, kept the structure but changed the factors:

1. Lack of control
2. Basic causes
3. Immediate causes
4. Incident
5. Loss or injury (Shahab Hosseinian & Jabbarani Torghabeh, 2012)

Bird's ideas first brought management into the equation, in the "lack of control" factor. The second domino, basic causes, was subdivided into two categories, of personal and job factors. Immediate causes was the new name given to "unsafe acts" in the original Domino Theory. Bird explained that immediate causes consisted of "substandard acts and conditions." They redefined incident as contact with "energy and substance." While it appears, to these authors, that little has changed from Heinrich to Bird, (besides the labels on the dominos) many modern safety experts utilize the Updated Domino Theory. According to Shahab, "The updated domino sequence can be used and applied to all types of accidents and is fundamental in loss control management" (Shahab Hosseinian & Jabbarani Torghabeh, 2012).

Many of the accidental models can trace their roots back to the pyramids. These models can be identified by their linear approach to accident causation. Perhaps one of the most dangerous ideas promoted by the pyramids is the idea "a multitude of minor incidents are required for one major incident to occur" (Awolusi & Marks, 2015). This is simply not true. Research shows that contributing factors for SIF and non-SIF incidents are not the same (Martin & Black, 2015). Prevention of papercuts, even 300 (See Figure 1) does not prevent a fatality, it likely doesn't prevent anything other than 300 papercuts.

The focus on eliminating these non-SIF injuries, has given rise to the safety profession's favorite number: zero.

Target zero

This erroneous concept that multiple minor incidents are necessary for major incidents to occur, has led to the wide acceptance of "target zero." Target zero, or goal zero initiatives are initiatives that promote the idea that *all* injuries and incidents can be prevented. This focus on eliminating all injuries, has led to an increase in SIF incidents (Martin & Black, 2015).

Target zero is a widespread challenge. The proponents of target zero, widely argue that any number other than zero is unethical. This is a logical fallacy known as a "false dilemma." No requirement for a goal exists! There is no evidence that supports that goal zero reduces SIF injuries, but rather evidence suggests the opposite, goal zero initiatives likely increase the risk of SIF incidents (Sherratt & Dainty, 2017). The focus on eliminating all injuries, rather than eliminating injuries that really matter has caused this increase (Sherratt & Dainty, 2017). At the very least, it can be shown that reduction of small injuries does not lead to a reduction in SIF potential injuries (Martin & Black, 2015).

Target zero, flies in the face of mathematical quantifications of risk. Safety has long struggled to implement concepts of risk management properly (Sanderson & Martin, 2022). Risk is often mathematically explained in the simple equation:

$$Risk = Probability \times Severity \qquad \text{(EQ 1)}$$

In basic risk management, loss control is divided into two broad categories. These are risk reduction (focusing on severity or consequence) and loss prevention (reducing loss frequency or probability). Since reducing the frequency (probability) or the consequence (severity) all the way to zero is typically not practicable, a small number will remain. This is the mathematical explanation for residual risk. Residual risk is defined as the risk that remains once loss control efforts are implemented. Target zero ignores the concept of "residual risks." This ignoring of residual risk, is often due to the low frequency of the problem. High severity events, with extremely low frequencies, can still represent a substantial risk mathematically. These substantial risks, are the safety embodiment of Taleb's Black Swan events (Taleb, 2010).

Opponents argue, that target zero is not a SMART goal. SMART, is an acrostic style mnemonic to help with goal setting. SMART goals should be:

- Specific
- Measurable
- Actionable
- Realistic
- Timed – that is identify specific timetable for events

Does target zero make a SMART goal? Target zero appears specific, but it falls apart in terms of SMART goals after that. Is it measurable? Proponents argue it is, after all, injuries are measured every day. But opponents argue that it is not measurable. How do you measure what did not happen? While it is easy to argue that something which occurred yesterday, but not today was reduced, what happens if it occurs tomorrow? Was the goal met? Is Target zero an actionable goal? What must a company do in order to achieve no injuries? Is it realistic? Can zero injuries actually be achieved? How long can zero injuries be sustained? All of these questions challenge the validity of target zero as a SMART goal.

Modern, as well as not so modern, research has given rise to arguments against target zero. In 1939, Dollard published his "frustration-aggression hypothesis" where he stated that:

> "…the occurrence of aggressive behavior always presupposes the existence of frustration and, contrariwise, that the existence of frustration always leads to some form of aggression" (Dollard et al., 1939)

Target zero is a source of frustration. Distribution of injuries is largely random (that is they cannot accurately be predicted) (Clemens, 2005). Random distributions of injuries lead to frustrated workforces, which in turn lead to aggressive behaviors (Dollard et al., 1939) (Breuer & Elson, 2017). These aggressive behaviors need not translate directly to violence, they often translate into extreme behavior to reach other goals (Breuer & Elson, 2017). For example, target zero initiatives are often reported with a decrease in injury reporting (Long, 2012) similar to other "safety incentive" programs (Wilkins, 2006).

Perhaps the biggest critique of target zero can be found in the UK construction industry (Sherratt & Dainty, 2017). Sherratt and Dainty write that:

> "you are actually marginally more likely to have major/specified accident working on a large UK construction site operated by a company that has adopted Zero within their safety strategy, than if you are working on a site without it.
>
> … Within the UK construction industry, there is no empirical evidence of greater improvements in safety management and resultant outcomes for companies that have adopted Zero than for those who have not. Zero has arguably stymied innovation and change, or at the very least maintained safety management practices to industry norms with the exception of Zero branding" (Sherratt & Dainty, 2017).

Here Dr. Sherratt confirms the idea that in safety, Black Swans (SIF injuries) hide in the mantra of target zero.

Problems with predictability

Much of occupational risk management involves the prediction of dangerous outcomes, commonly referred to as hazard identification. Practitioners have learned, or been taught, to identify activities that present the highest risks, either in terms of frequency (probability of loss) or severity (gravity of consequence). They make the predictions, but are the predictions they make valid? Philip Tetlock, has studied this question extensively: "Do experts make better predictions?" (Lerner & Tetlock, 1999; Tetlock, 2002; Tetlock & Kim, 1987). Tetlock identified that experts rarely outperformed basic algorithms, and often underperformed them (Taleb, 2010; Tetlock, 1999).

Tetlock explained that experts are often prisoners of perception (Tetlock, 1999). Tetlock explains that experts tend to see their incorrect predictions as "almost right" rather than "partially wrong" (Tetlock, 1999). While this may seem simple semantics, it creates a problem. Experts don't know when they are wrong. Taleb further explains several reasons why experts fail to produce better predictions. He offers several reasons:

- They convince themselves they were playing a different game
- They invoke the principle of the outlier
- They tell themselves they were “almost right”
- They make the claim it is not their job to know everything (Taleb, 2010)

Experts tend to not be any better at predicting the future than others. All of this leads to the following questions: If risk reporting works, why do we under forecast serious injuries and fatalities? If risk reporting does not work, why do we do it at all?

This error in prediction results in the improper classification of risks. Small risks are often overestimated, while large risks are frequently underestimated (Hubbard, 2020). Humans are intrinsically poor judges of low frequency events. We do not inherently understand them. These misjudgments often result in a focus on small events due to their higher probabilities. This leads to a general neglect of large severity, due to their extremely low frequency (Taleb, 2010). Often, this results in a decrease of overall injuries, but an increase in SIF injuries.

Additionally, these errors are compounded by what Taleb refers to as the two realms of “Extremistan” and “Mediocristan” (Taleb, 2010). In Mediocristan, the law of large numbers applies. In other words, no single datum point will significantly skew the average too drastically (Taleb, 2010). Examples of variables in the Mediocristan realm include things like height, weight, age, etc. On the other hand, variables in the Extremistan realm would be extremely skewed by a single datum point. Examples of variables in Extremistan would include: income, net worth, numbers of books sold by authors, and others (Taleb, 2010). Most of the prediction tools that are used in practice are designed for Mediocristan, and they perform well within that realm. When they are applied to Extremistan, they fail, and they fail spectacularly (Taleb, 2010).

Common practice in risk management, and safety, is to review losses. Losses are reviewed, analyzed, trended, studied, and used as predictors for future losses. These are methods from Mediocristan, they in no way can predict a new event (Taleb, 2010). These predictions create a false sense of security, that what *did not happen* yesterday, *cannot happen* tomorrow. Safety Black Swan events thrive in this contrived notion of security.

Inherent in these data that are crunched and analyzed are biases. Routine tasks, generate the most data, and therefore are the most closely scrutinized. The tasks that are best known, continue to be the most studied. Rarely are non-routine tasks examined. SIF’s largely occur in non-routine tasks. Prediction in Extremistan is not done by looking at evidence, evidence often comes too late to be of value, rather prediction in Extremistan should be done by looking at the potential for harm (Taleb, 2010). Put practically, rather than examine a non-routine task for the harm that is has caused previously, it should be examined for the potential harm that could occur should something go wrong.

Because prediction based of previous results is flawed in Extremistan, something new needs to be done. Requiring robust job hazard analysis, or pre-task analysis, can help better identify hazards in Extremistan, than looking at yesterday’s failures to predict large losses.

Incident investigations – creating Black Swans from past events

Problematic predictions run both ways. This is something that is often forgotten or ignored in the safety industry. It is often believed that incident investigations are based on facts, and not a form of prediction. Consider the following thought experiment posed by Taleb:

> "Operation 1 (the melting ice cube): Imagine an ice cube and consider how it may melt over the next tow hours. Try to envision the resulting puddle.
>
> Operation 2 (where did the water come from?): Consider a puddle of water on the floor. Now try to reconstruct in your mind's eye the shape of the ice cube it may have once been. Note that the puddle may not necessarily have originated from an ice cube." (Taleb, 2010)

When reviewing the past, we have to make assumptions about where things began. These assumptions, are subject to the same biases, and challenges that forward projecting predictions face, and the problem becomes exponentially more difficult (Taleb, 2010). This inherent challenge to predict the past renders many incident investigations ineffectual.

If we are unable to identify the shape of an ice cube from a puddle, when we are able to predict the shape of a puddle from an ice cube, how are we going to accurately predict past behaviors? The simple answer is: we cannot. We cannot even accurately predict behavior in the forward direction. In order to simplify this, simple linear causation models are utilized. Imagine that each known event is a simple dot on a page. The assumption is that the dots are connected in a linear relationship, that there is only a single straight line that connects multiple dots. However, if the dots are fitted with a curve, there is an infinite number of curves that can connect all of the points (Taleb, 2010). This assumption of linear causality is manifested plainly in the popular "Five Why's" method. This method artificially imposes a single cause to each step in the procedure. The five whys method is ill equipped for determining multiple causal factors. Each successive "why" forces the investigator closer to a single cause.

A single cause is a great concept, but it does not reflect reality, the world simply doesn't operate in such a linear fashion (Taleb, 2016). Fred Manuele writes:

> "Many incidents resulting in serious injury or fatality are unique and singular events, having multiple and complex causal factors that may have organizational, technical, operational systems, or cultural origins" (F. A. Manuele, 2014)

This lack of linearity, is a major flaw with James Reason's Swiss Cheese Model. Events don't occur in nice tidy orders, events in the real world are not subject to those rules of the game. Taleb has termed this assumption that the world follows the specific rules (like a game) the "Ludic Fallacy" (Taleb, 2010).

Probability theory as it generally known today, was invented by Blaise Pascal and Pierre de Fermat in the 1650's (Bernstein, 1996). It was invented as a way to compute the odds in games of chance (gambling) and is the basis for much of modern risk management (Lyon & Popov, 2021). The Ludic fallacy brings to light a world that does not follow the same rules as games of chance (Taleb, 2010).

The Ludic fallacy causes the oversimplification of the real world. This oversimplification causes operational blind-spots to appear. Safety professionals are blinded by what has happened in the past, unwilling to consider emerging risks simply because they have not occurred previously. This allows SIF incidents to occur, specifically because they often occur in non-routine tasks (Martin & Black, 2015).

Furthermore, while incident investigations are fundamental in most safety management systems, research shows that most incident investigations, and the methods upon which they are based, are flawed (F. A. Manuele, 2014). These flawed methods often lead to a (single) root cause of human error (Hollnagel & Amalberti, 2001). If the object of an incident investigation is to truly find a root cause, human error is never an acceptable end point. This is discussed more fully in the following section.

A popular concept that has risen, related to incident investigations is the concept of a "near miss." Near misses, sometimes referred to as "near hits," "close calls," or "zero cost losses," are deviations from normal operations, that do not result in a loss to property or injuries. There has been a push to include near misses, near miss reporting, and near miss investigations as a type of leading indicator for a safety system. But learning from the past is difficult, especially in cases of near misses (Dillon & Tinsley, 2008).

Near misses, provide more information about the system in action. If Bayesian methods are applied, this should lead to an increase in predictive powers, and therefore an increase in safety performance. Alas, in reality it does not (Dillon & Tinsley, 2008). Near misses, in spite of the attention, are internally viewed as successes, precisely because of the "no harm, no foul" ideology. This gave rise to the "near-miss interpretation paradox" which states:

> "People with near miss information were *more* likely to make the riskier choice than people who did not have information about near misses" (Dillon & Tinsley, 2008; Hubbard, 2020).

Investigating near misses, provides more information, and is likely to result in riskier decisions being made by those with the information. In a colloquial expression, "if it ain't broke don't fix it" one can see the near-miss interpretation paradox. Afterall, the system functioned as intended, no loss occurred, why make a change? This is precisely the reason why near-misses are investigated, to identify errors and correct them before a loss occurs. According to research, the observed behavior is drastically different from the desired behavior when it comes to near miss information (Dillon & Tinsley, 2008).

Problems with automation

The "New View" of safety, sometimes called "Safety II" or "Safety Differently" embraces the idea that human error is an inescapable part of life (Conklin, 2019). As an inescapable part of life, it should never come as a surprise. Safety Differently (Dekker, 2014), and human and organizational performance (Conklin, 2019) encourage systems to anticipate human error. This approach, may also lead to SIF incidents. This is counterintuitive to what many believe, but it was aptly explained by Earl Wiener in what would be coined "Wiener's Laws" (Harford, 2016). One of his laws: "Digital devices tune out small errors, while creating opportunities for large ones" (Langewiesche, 2014).

But how does this happen? As systems are designed to anticipate human error, they move closer to automation. Automation can be a good thing, it can reduce costs and reduce injuries. The problem is that automated systems follow programs exactly as they are told. They do not account for changes in the operating conditions. Thus, when the system is overwhelmed, or cannot function within its given parameters, control is passed over to the human operator. This means that the operator, who has been unable to practice their skills with normal operations, is taking over at the most extreme time (Harford, 2016; Van Cott et al., 1996). This is the definition of a non-routine task, and leads to SIF incidents.

This is known as the "automation paradox" and is well documented. It has been observed with operators of self-driving vehicles (for example an Uber driver), to airline pilots (Air France Flight 447), even to school kids (being unable to perform basic arithmetic without a calculator). When the stakes are highest, automation has essentially relegated humans to putting in a bench player. Gary Klein summarized this idea best:

> "When the algorithms are making the decisions, people often stop working to get better. The algorithms can make it hard to diagnose reasons for failures. As people become more dependent on algorithms, their judgement may erode, making them depend even more on the algorithms. That process sets up a vicious cycle. People get passive and less vigilant when algorithms make the decisions" (Klein, 2011).

Humans are not allowed to operate the system regularly, but are only allowed to operate the system in extreme circumstances, a recipe for Black Swans in the safety world.

Procrustes bed

In Greek mythology, Procrustes was a demigod who preyed on travelers in Attica. He would convince the travelers, to lie on a bed made of iron, but Procrustes insisted that they fit the bed perfectly. He would stretch their limbs to make them fit if they were too short, or cut off a section of the traveler's legs if they were too tall. Taleb utilizes this myth in his term "Procrustean bed" (Taleb, 2016).

A Procrustean bed is an example of solving a problem by focusing on the wrong variable. In SIF prevention, Procrustean beds abound. Taleb writes:

> "…treating an organism like a simple machine is a kind of simplification or approximation or reduction that is exactly like a Procrustean bed. It is often with the most noble intentions that we do so, as we are pressured to 'fix' things, so we often blow them up with our fear of randomness and our love of smoothness" (Taleb, 2016).

Most modern safety theories fall into this particular Procrustean bed. Behavior-based safety (BBS) tends to reduce human behavior to a simple 3 step system, the ABC's of behavior, Antecedent, Behavior, Consequence. The crux of the theory is that humans are influenced by antecedents, but motivated by consequences (Geller, 2005). This reduces human behavior, a complex and unpredictable phenomenon, to a simple predictive model.

That is not to say that behavior-based safety is bad, or wrong, but simply incomplete. The more modern theory of human and organizational performance (HOP) also reduces human behavior to a simple predictive model. Rather than focus on the behavior being driven by the consequence, human and organizational performance focus on the surroundings and context of the behavior (Conklin, 2019). As mentioned previously, the idea is that human error should be accounted for, and designed out of the system. This puts a large focus on controlling the antecedent in ABC model. In 2017, Conklin and Geller were part of a Safety plenary panel for the American Society of Safety Engineers virtual classroom (ASSE SafetyVC, 2017) where they debated the merits of each of the theories they were promoting. Conklin argues that behavior is an output of a system, Geller argues that behavior is an input in the system. The debate over behavior in this context is a Procrustean bed.

Focusing on this particular Procrustean bed, has caused a lot of debate in the safety profession. Professionals will loudly and proudly defend their particular brand of safety practice. Opponents of BBS will argue that it "blames the worker" for the incident. Opponents of HOP will argue with equal fervor by sarcastically remarking "humans are the solution" to any incident where human factors have been causal. Both theories are focusing on improving human choices, but both are approaching behavior from different sides of the aisle. Both sides shout at the other that they simply "misunderstood" or "misapplied" all the while reassuring themselves that they have properly understood the other side.

Yet the fact remains, that SIF incidents have occurred (and continue to occur) in organizations that practice BBS, as well as organizations that practice HOP. The increased focus from either theory, falls into the same routines, focuses on the same reduction of all injuries, and "increasing safety" as though that has any meaning. Choosing a safety theory to which one subscribes appears to be Procrustean bed.

The very problem of SIF incidents tends to force a Procrustean bed. The natural tendency is to focus on loss prevention (reducing the frequency of incidents) as a means of loss control, but SIF incidents tend to be rare, but high in severity. A loss reduction (reducing the severity of incidents) is what is needed for SIF reduction. Again, this presents an argument that is difficult to swallow for many safety professionals. If an incident never occurs, it will have a severity of zero. The challenge lies in the achievement of zero, eventually the residual risk gets enough exposure to occur. When dealing with large consequences, such as SIF incidents, it is best to focus on the consequences (which can be understood) rather than on probabilities, which are often misunderstood (Taleb, 2010).

A final challenge

Perhaps the largest challenge for SIF incidents, is our failure to learn from them. SIF incidents often hide in plain sight. Martin and Black write:

> "All recordable injuries are not equal. A broken foot caused by stepping on a rock in the parking lot has significantly less SIF exposure than a broken foot that is the result of being driven over by a forklift. On the OSHA 300 log, these two cases appear identical due to outcome, but the exposure situation tells a different story" (Martin & Black, 2015).

SIFs need to be examined at an exposure level, based on the consequence that may occur, not on the probability that it may never occur. As companies continue to utilize arbitrary risk matrices and non-quantified risk calculations, the SIF incidents will continue to occur (Sanderson & Martin, 2022).

Failure to learn from SIF incidents also results from the investigations into SIF incidents. Investigations following SIF injuries should be immediate and thorough, and should be prioritized higher than a typical incident investigation (F. A. Manuele, 2014). Yet, often SIF injuries are investigated half-heartedly. They are done quickly, because the demand for the information is immediate. Investigators know that the executives, or board members will be reviewing the report. Add to this that the investigator is often the front-line supervisor, and you have a recipe for incomplete information (F. A. Manuele, 2014).

Not only are reports often rushed, but they generally reinforce the belief that "had the person acted differently, the outcome would have been different" (ASSE SafetyVC, 2017). This seems to reinforce the "I knew that all along" bias in many organizations (Taleb, 2010). Reports are created not for learning, but for an exercise in covering one's backside.

Conclusion

Black Swans in safety, or SIF incidents, continue to be a challenge for the profession. In spite of numerous articles addressing them, and identifying that they do not have the same precursors or causal factors as non-SIF incidents, they continue to be an afterthought in the safety profession (F. A. Manuele, 2014; Martin & Black, 2015). The continued push for elimination of all harm, has led to a significant oversight of serious harm (Sherratt & Dainty, 2017).

Many of the safety theories, and practice disciplines are increasing the potential for SIF incidents (Martin & Black, 2015). New approaches, new paradigms, and new methods have been proposed to reduce SIF incidents, but those continue to be neglected topics in the research. Follow up is needed to verify whether or not these proposed methods, paradigms, or approaches will have an impact.

SIF incidents, seemingly appear to be random, due to the inability to predict them accurately. Taleb writes "In practice, randomness is fundamentally incomplete information" (Taleb, 2010). Safety professionals, need to understand that with SIF incidents, the focus is on the severity or the consequence, not the likelihood of occurrence. This too will require a paradigm shift. Current safety practice focuses on a hierarchy of controls, elimination > substitution > engineering controls > administrative controls > personal protective equipment. The most preferred controls focus on loss prevention (frequency), but when it comes to SIF incidents, a focus needs to be on loss reduction (severity).

References

1. ASSE SafetyVC. (2017, July 11). *Safety 2017 - Plenary Town Hall - YouTube*. https://www.youtube.com/watch?v=43u61TQpOxQ&ab_channel=SafetyVC
2. Aven, T. (2015). Implications of black swans to the foundations and practice of risk assessment and management. *Reliability Engineering & System Safety*, *134*, 83–91. https://doi.org/10.1016/J.RESS.2014.10.004
3. Awolusi, I., & Marks, E. (2015). *Near Miss Reporting to Enhance Safety in the Steel Industry Wearable Sensing Devices and the Internet of Things for Worker Safety and Health Management in the Steel Manufacturing Industry View project Pro-Active Safety Technologies and Innovation for Manufacturing Environments View project Near Miss Reporting to Enhance Safety in the Steel Industry*. https://www.researchgate.net/publication/320345534

4. Bernstein, P. L. (1996). *Against the gods: The remarkable story of risk*. John Wiley & Sons, Ltd.
5. Bird, F. E., & Germain, G. L. (1996). *Practical loss control leadership*. Det Norske Veritas (USA).
6. Breuer, J., & Elson, M. (2017). Frustration-Aggression Theory. *The Wiley Handbook of Violence and Aggression*, 1–12. https://doi.org/10.1002/9781119057574.WHBVA040
7. Canfora, C., Ottmann, A., Gutenberg-Die, J., & Mannheim, R. (. (2018). *Of Ostriches, Pyramids, and Swiss Cheese-Risks in Safety-Critical Translations*. https://www.researchgate.net/publication/326317704
8. Chute, R. D., & Wiener, E. L. (1995). Cockpit-Cabin Communication: I. A Tale of Two Cultures. *The International Journal of Aviation Psychology*, *5*(3), 257–276. https://doi.org/10.1207/s15327108ijap0503_2
9. Clemens, P. (2005). Zero-Injury Workdays. *Professional Safety*, *May*, 40–41. https://aeasseincludes.assp.org/professionalsafety/pastissues/050/05/040505as.pdf
10. Conklin, T. (2019). *The 5 Principles of Human Performance: A contemporary update of the building blocks of Human Performance for the new view of safety*. Pre-Accident Investigation Media.
11. Dekker, S. (2014). *Safety differently*. CRC Press London.
12. Dillon, R. L., & Tinsley, C. H. (2008). How near-misses influence decision making under risk: A missed opportunity for learning. *Management Science*, *54*(8), 1425–1440. https://doi.org/10.1287/mnsc.1080.0869
13. Dollard, J., Miller, N., Doob, L., Mowrer, O., & Sears, R. (1939). *Frustration and aggression.* https://psycnet.apa.org/record/2004-16227-000
14. Gallivan, S., Taxis, K., Franklin, B. D., & Barber, N. (2008). Is the principle of a stable Heinrich ratio a myth? A multimethod analysis. *Drug Safety*, *31*(8), 637–642. https://doi.org/10.2165/00002018-200831080-00001
15. Geller, E. S. (2005). Behavior-based safety and occupational risk management. *Behavior Modification*, *29*(3), 539–561. https://doi.org/10.1177/0145445504273287
16. Harford, T. (2016). *Messy: The Power of Disorder to Transform Our Lives*. https://books.google.com/books?hl=en&lr=&id=P9gzDwAAQBAJ&oi=fnd&pg=PA1&dq=Tim+Harford+Messy&ots=HtBX68OBTH&sig=eAZRkEKCKTt2GDqzZqsF4mu5rOQ
17. Heinrich, H. (1931). *Industrial accident prevention; a scientific approach.* https://psycnet.apa.org/record/1931-10417-000
18. Hollnagel, E., & Amalberti, R. (2001). The emperor's new clothes: Or whatever happened to "human error." *Procedings of the 4th International Workshop on Human Error, Safety Systems, and Development*, 1–18.
19. Hubbard, D. W. (2020). *The Failure of Risk Management: Why It's Broken and How to Fix It*. John Wiley & Sons, Ltd. https://books.google.com/books?hl=en&lr=&id=fMbKDwAAQBAJ&oi=fnd&pg=PR11&dq=Failure+of+Risk+Management&ots=khabyUfWoY&sig=1rhKBAsqdD8PaI9Hkwiboj_lc5E
20. Klein, G. A. (2011). *Streetlights and shadows: Searching for the keys to adaptive decision making*. MIT Press.
21. Langewiesche, W. (2014). The human factor. *Vanity Fair*, *17*.
22. Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. In *Psychological Bulletin* (Vol. 125, Issue 2, pp. 255–275). https://doi.org/10.1037/0033-2909.125.2.255
23. Long, R. (2012). *For the Love of Zero : Human Fallibility and Risk.* Scotoma Press.
24. Lyon, B. K., & Popov, G. (2021). *Assessing and Managing Risk An ERM Perspective*. American Society of Safety Professionals.
25. Manuele, F. (2003). *On the practice of safety*. https://books.google.com/books?hl=en&lr=&id=H4ou_64UL6MC&oi=fnd&pg=PR7&dq=F.A.+Manuele&ots=bDU3MhWvYR&sig=HehUQrfMDTy7mr99owLjItqFS8M
26. Manuele, F. A. (2014). Incident Investigation Our Methods Are Flawed. *Professional Safety*, *October*, 37–43. https://doi.org/10.1016/j.jsr.2008.02.019
27. Martin, D., & Black, A. (2015). Preventing Serious Injuries and Fatalities: Study Reveals Precursors and Paradigms. *Professional Safety*, *60*(09), 35–43. https://onepetro.org/PS/article-abstract/60/09/35/33319
28. Sanderson, M. (Capitol T. U., & Martin, L. F. (2022). *Challenges of risk management in safety*

[Submitted for publication].

29. Shahab Hosseinian, S., & Jabbarani Torghabeh, Z. (2012). MAJOR THEORIES OF CONSTRUCTION ACCIDENT CAUSATION MODELS: A LITERATURE REVIEW. *International Journal of Advances in Engineering & Technology*, *4*(2), 53–66.
30. Sherratt, F., & Dainty, A. R. J. (2017). UK construction safety: A zero paradox? *Policy and Practice in Health and Safety*, *15*(2), 108–116. https://doi.org/10.1080/14773996.2017.1305040
31. Taleb, N. N. (2010). *The black swan: The impact of the highly improbable*. Random House Trade Paperbacks.
32. Taleb, N. N. (2016). *Antifragile: Things that Gain from Disorder*. Random House Trade Paperbacks. https://doi.org/10.1177/1755738019885153
33. Tetlock, P. E. (1999). Theory-Driven Reasoning About Plausible Pasts and Probable Futures in World Politics: Are We Prisoners of Our Preconceptions? *American Journal of Political Science*, *43*(2), 335. https://doi.org/10.2307/2991798
34. Tetlock, P. E. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review*, *109*(3), 451–471. https://doi.org/10.1037/0033-295X.109.3.451
35. Tetlock, P. E., & Kim, J. I. (1987). Accountability and Judgment Processes in a Personality Prediction Task. *Journal of Personality and Social Psychology*, *52*(4), 700–709. https://doi.org/10.1037/0022-3514.52.4.700
36. Van Cott, H. P., Wiener, E. L., Wickens, C. D., Blackman, H. S., & Sheridan, T. B. (1996). Smart automation enhances safety: A motion for debate. *Ergonomics in Design*, *4*(10), 19. https://doi.org/10.1177/154193129503900901
37. Wilkins, S. L. (2006). Safety Incentive Programs: A Case Study. *Unpublished*.